Porting to Columbia: With SGI's MPT

Category: Porting to Columbia

DRAFT

This article is being reviewed for completeness and technical accuracy.

The available SGI MPT modules on Columbia are:

```
mpt.1.16.0.0
mpt.1.18.0.0
mpt.1.19.0.0
mpt.1.22.0.0
mpt.1.25
```

The current default version is *mpt.1.16.0.0*.

Environment Variables

On Columbia, when you load any of the above MPT modules, several environment variables such as CPATH, INCLUDE, LD_LIBRARY_PATH, etc., are modified by pre-pending the appropriate MPT directories. Also, the following MPT-related environment variables are modified from their default values for improved performance:

```
setenv MPI_BUFS_PER_HOST 256
setenv MPI_BUFS_PER_PROC 256
setenv MPI_DSM_DISTRIBUTE
```

The meanings of these variables and their default values are:

- MPI_BUFS_PER_HOST

  Determines the number of shared message buffers (16 KB each) that MPI is to allocate for each host (i.e., C21, C22, C23, C24). These buffers are used to send and receive long inter-host messages.

  Default: 32 pages (1 page = 16KB) for mpt.1.16, mpt.1.18, mpt.1.19, mpt.1.22
  Default: 96 pages (a page = 16KB) for mpt.1.25

- MPI_BUFS_PER_PROC

Determines the number of private message buffers (16 KB each) that MPI is to allocate for each process. These buffers are used to send long messages and intra-host messages.

Default: 32 pages (1 page = 16KB)

- MPI_DSM_DISTRIBUTE (toggle)

  Activates NUMA job placement mode. This mode ensures that each MPI process gets a unique CPU and physical memory on the host with which that CPU is associated. This feature can also be overridden by using dplace or omplace. This feature is most useful if running on a dedicated system or running within a cpuset.

  Default: Not enabled

Building Applications

Building MPI applications with SGI's MPT library simply requires linking with -lmpi and/or -lmpi++. See the article SGI MPT for some examples.

Running Applications

```
MPI executables built with SGI's MPT are not allowed to run on the
Columbia front-end node.
```

You can run your MPI job on C21 - C24 in an interactive PBS session or through a PBS batch job. Use **mpiexec** (under /PBS/bin) or mpirun to start your MPI processes. For example:

```
#PBS -lncpus=8
....
mpiexec -np N ./your_executable
```

The -np flag (with *N* MPI processes) can be omitted if the value of *N* is the same as the value specified for *ncpus*.

Article ID: 107
Last updated: 08 Mar, 2011
Computing at NAS -> Porting & Developing Applications -> Porting to Columbia -> Porting to Columbia: With SGI's MPT
http://www.nas.nasa.gov/hecc/support/kb/entry/107/?ajax=1